# Blind Estimation of HDR Image Quality

Navaneeth K. K.

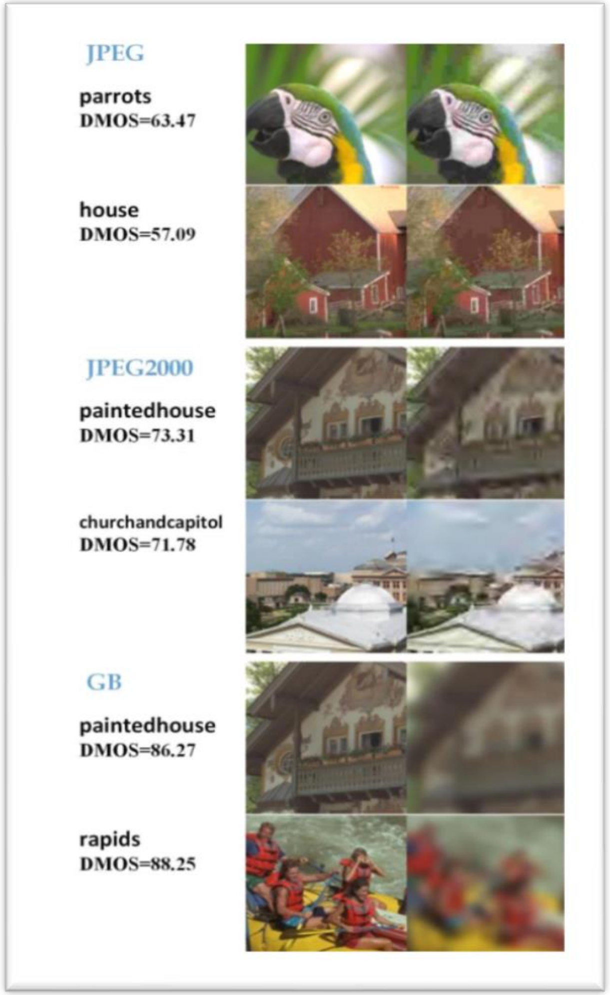
PhD, Multimedia Research Centre

# Contents

- Introduction to concepts
- Problem and relevance
- Proposed solution
- Results
- Conclusion

# Image Quality Assessment

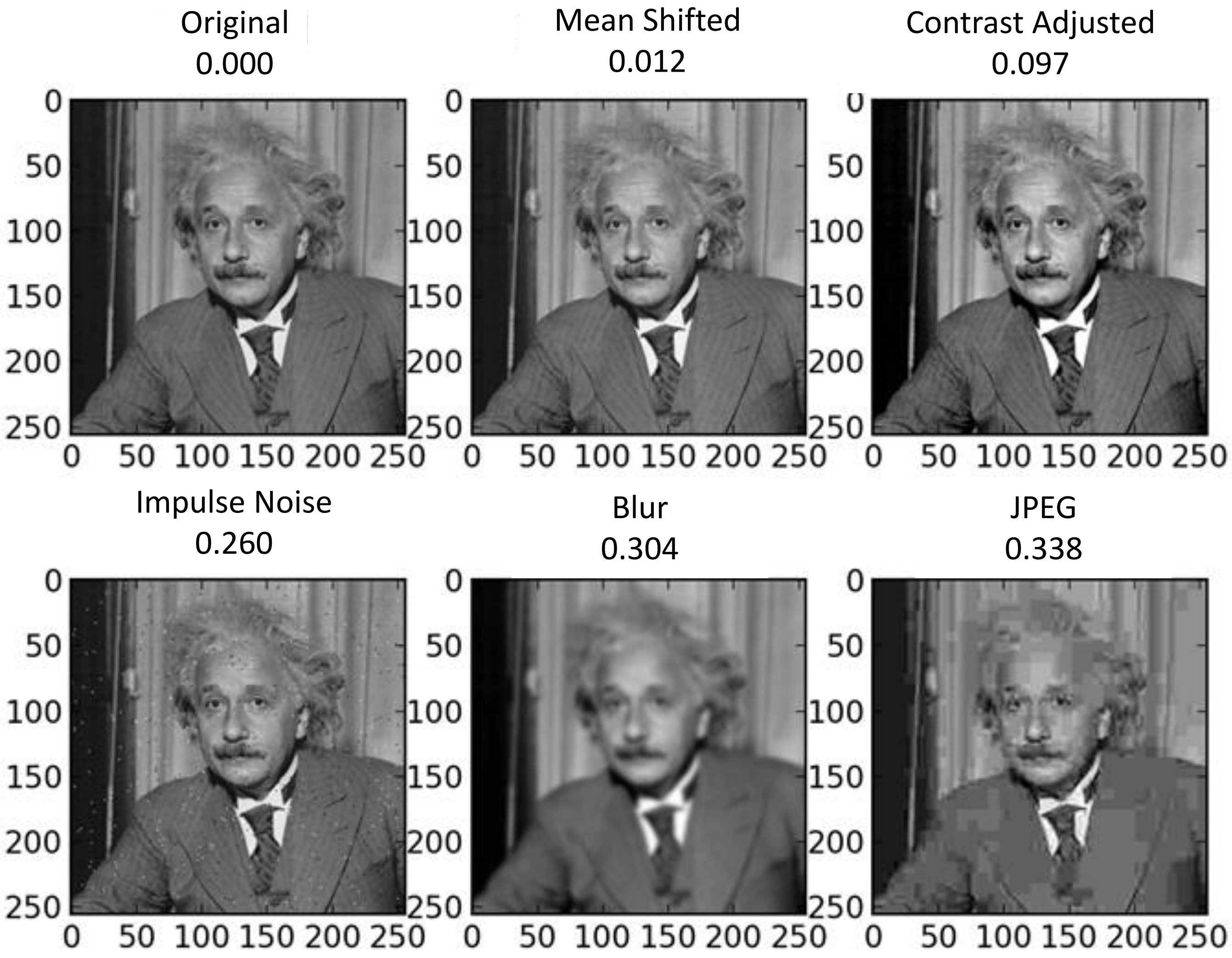
- Objective measure of perceived quality of image.
- Psycho-visual experiments to determine the limits of HVS

# Objective Quality Assessment

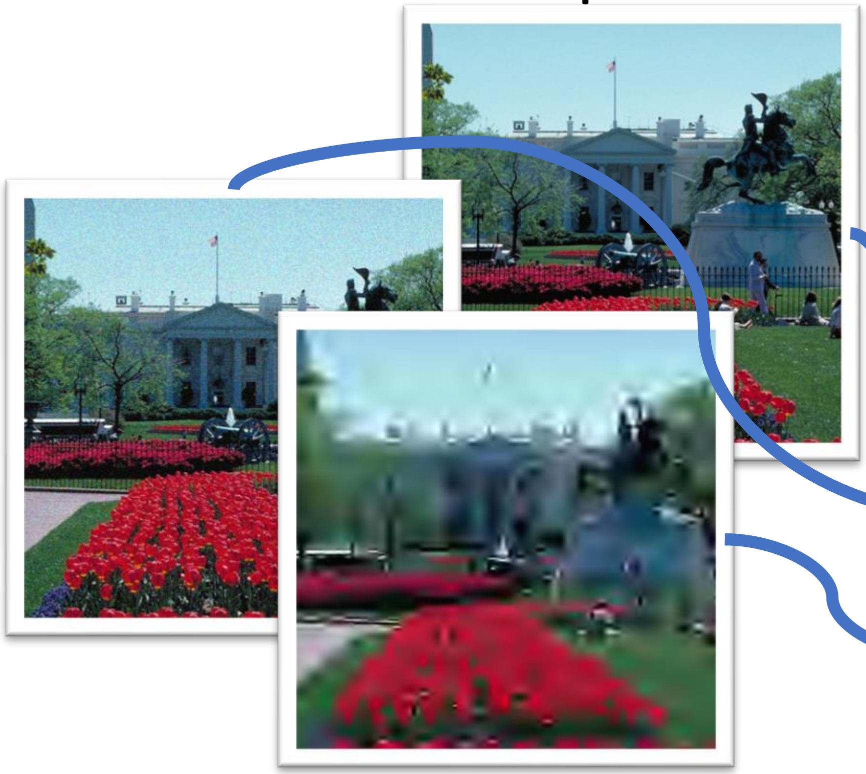
- DMOS – Difference in Mean Option Scores
- Large DMOS -> Very noisy
- Is measure of Perceived noise.

- **Datasets :**
  **Clean-noisy image pair, DMOS for noisy image.**



# Quality Assessment Algorithms

- Automated algorithms that can predict a quality score for content that can correspond to human judgement of quality.



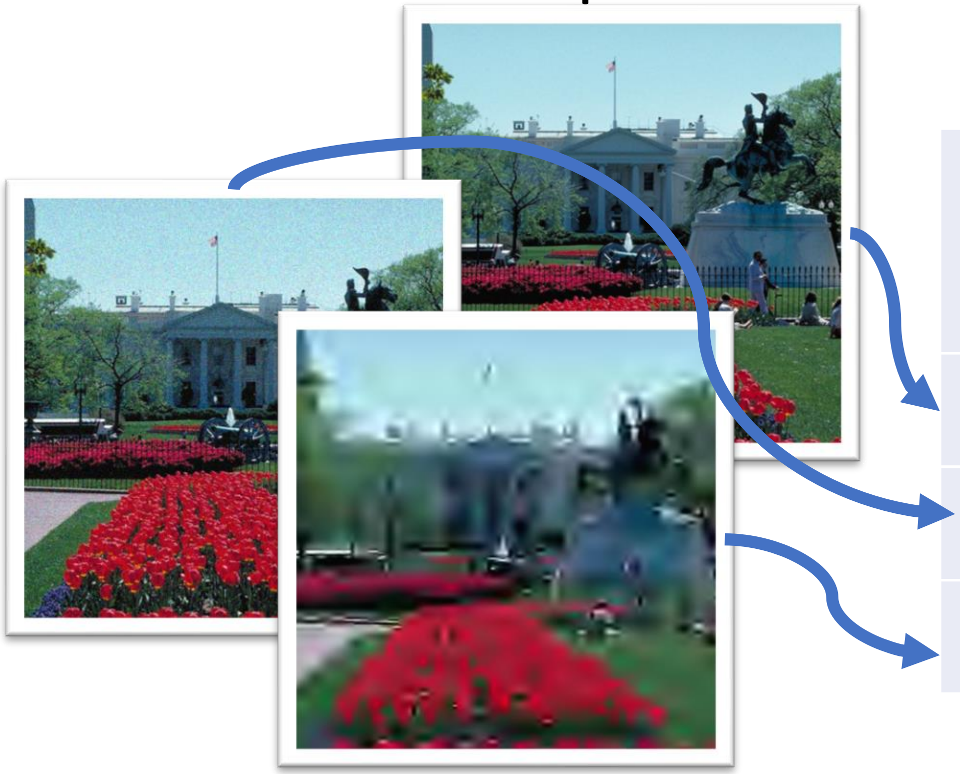
| Human score | Algorithm 1 | Algorithm 2 |
|---|---|---|
| 5 | 200 | 1 |

Measured by correlation scores. Popular metrics are
Spearman correlation coefficients (**SRCC**)
Pearson correlation coefficients (**PLCC**)

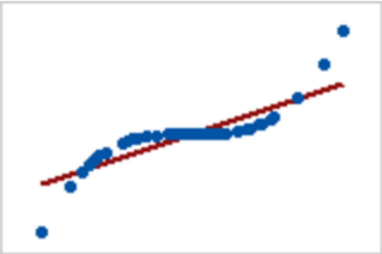
Higher = Better

# Quality Assessment Algorithms

- Automated algorithms that can predict a quality score for content that can correspond to human judgement of quality.



| Human sc | rithm 2 |
|---|---|
| 5 | 1 |



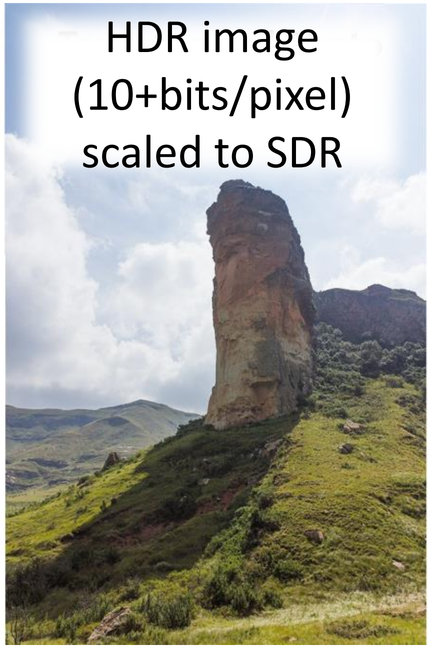
Pearson = +0.851, Spearman = +1

Measured by correlation scores. Popular metrics are
Spearman correlation coefficients (**SRCC**)
Pearson correlation coefficients (**PLCC**)

Higher = Better

# High Dynamic Range Imaging

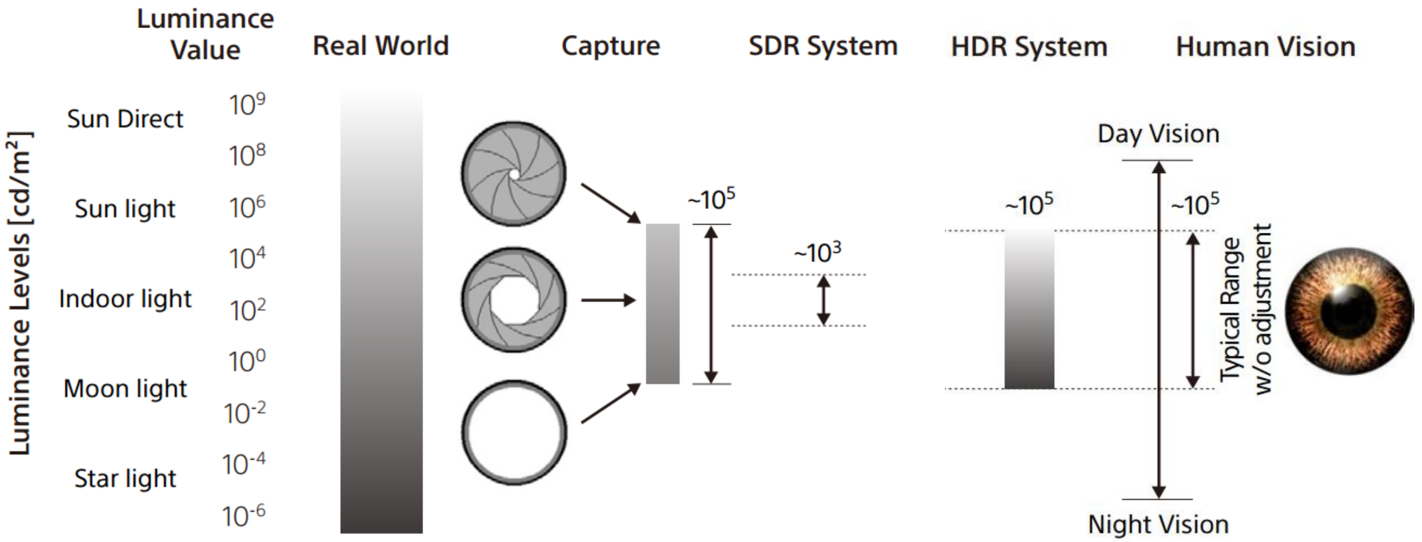
- Use of larger number of bits in imaging to display a larger range of luminance (brightness) values.


Courtesy: http://adaptivesamples.com/2016/03/16/make-your-own-hdri/

# High Dynamic Range Imaging

- It all the technologies involved in use of 10+ bit to encode images *as opposed to conventional 8 bits in Low Dynamic range images.*
- The compatible displays need  to display large luminance range.




Unit of light intensity: $Cd/m^2$

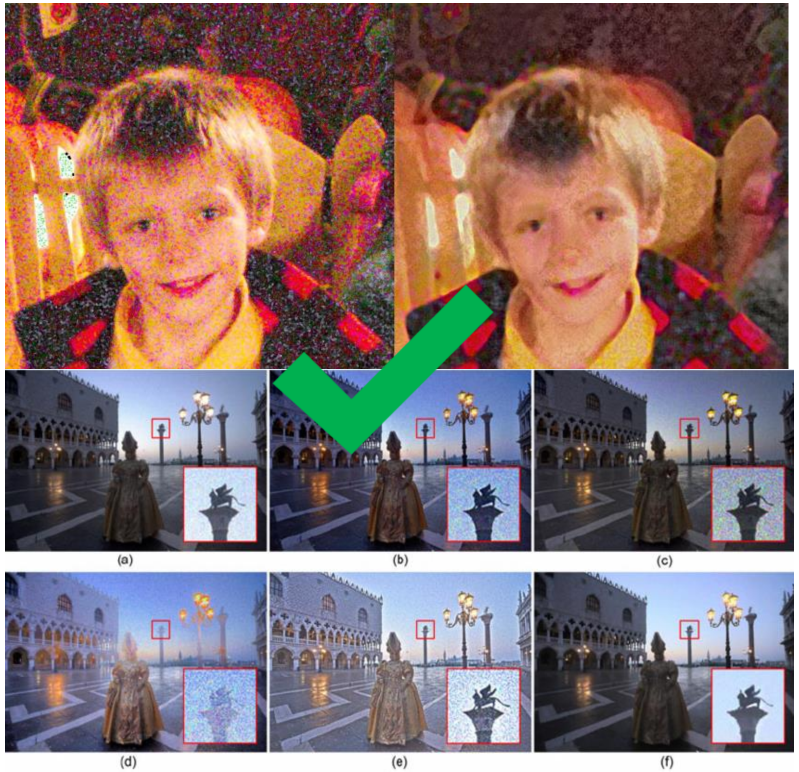
Courtesy: https://pro.sony.com/bbsccms/assets/files/cat/hdr/latest/MK20109V1_1_HDR.pdf

# Problem

- Design of Image quality assessment algorithm for High Dynamic Range Images.

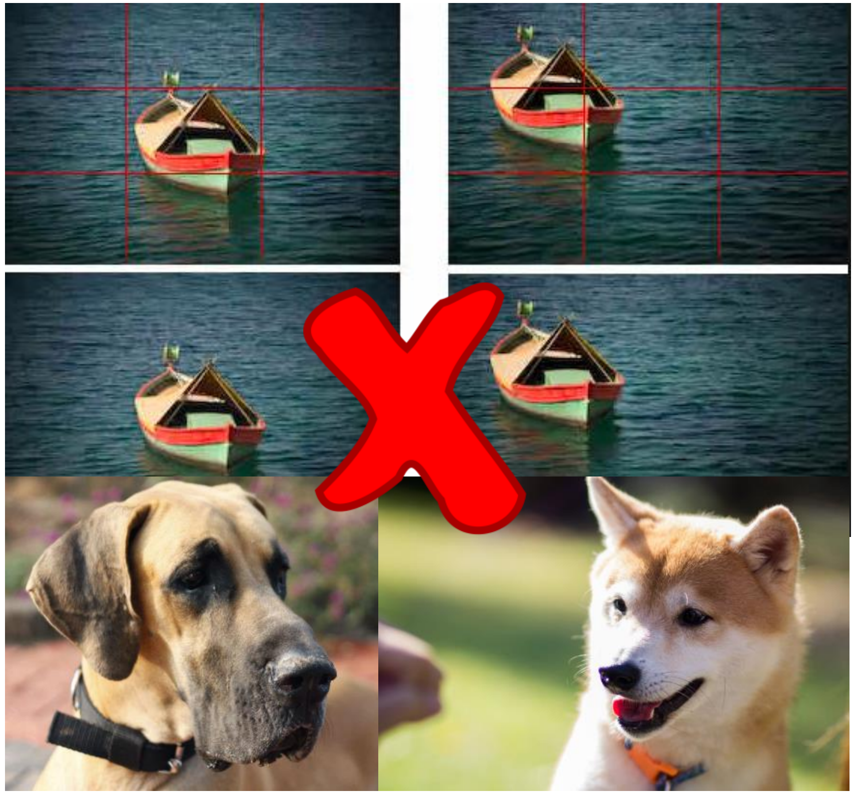
- *This research is limited to Low level distortions and its effects on quality.*

- *Noise, Image compression artifacts, changes in contrast etc are modelled.*



# Problem

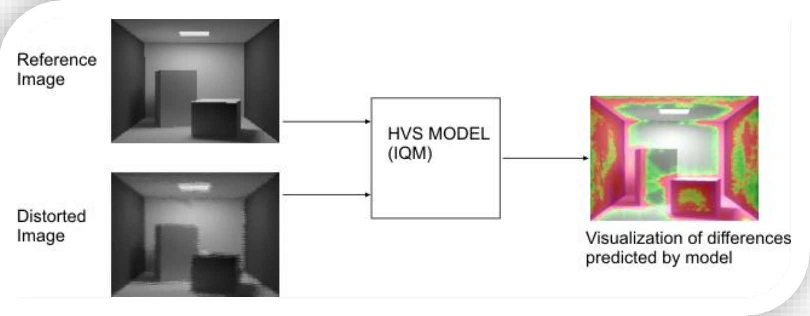
- Design of Image quality assessment algorithm for High Dynamic Range Images.

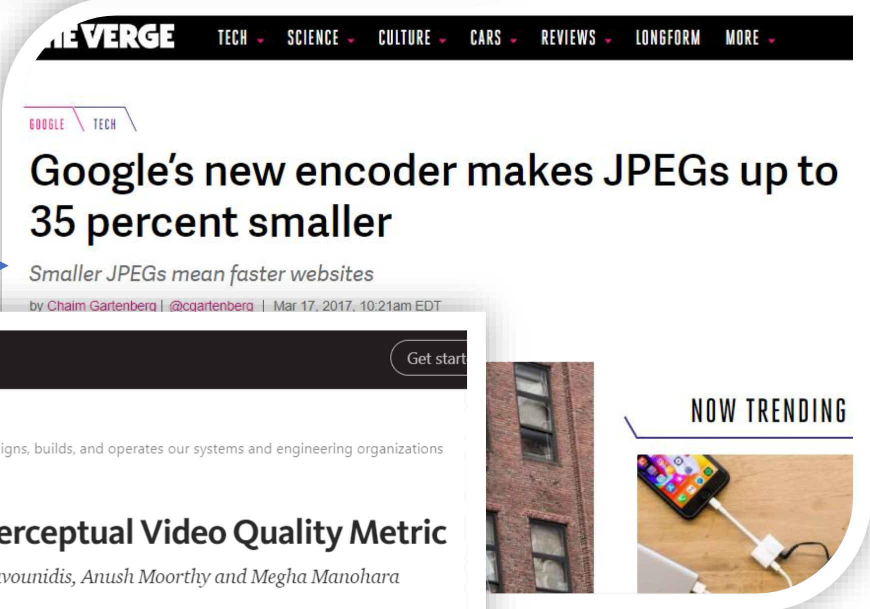


- *This research is limited to Low level distortions and its effects on quality.*
- *Noise, Image compression artifacts, changes in contrast etc are modelled.*

# Why

- Works behind the scenes in a lot of fields.
  - **Better displays.**
  - **Improved algorithms.**
  - **Efficient compression of image.**
  - **Better video codec.**
  - **Comparing algorithms.**





THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM MORE

GOOGLE TECH

## Google's new encoder makes JPEGs up to 35 percent smaller

*Smaller JPEGs mean faster websites*

by Chaim Gartenberg | @cgartenberg | Mar 17, 2017, 10:21am EDT

NOW TRENDING


THE NETFLIX TECH BLOG

Netflix Technology Blog

Learn more about how Netflix designs, builds, and operates our systems and engineering organizations

Jun 5, 2016 · 19 min read

## Toward A Practical Perceptual Video Quality Metric

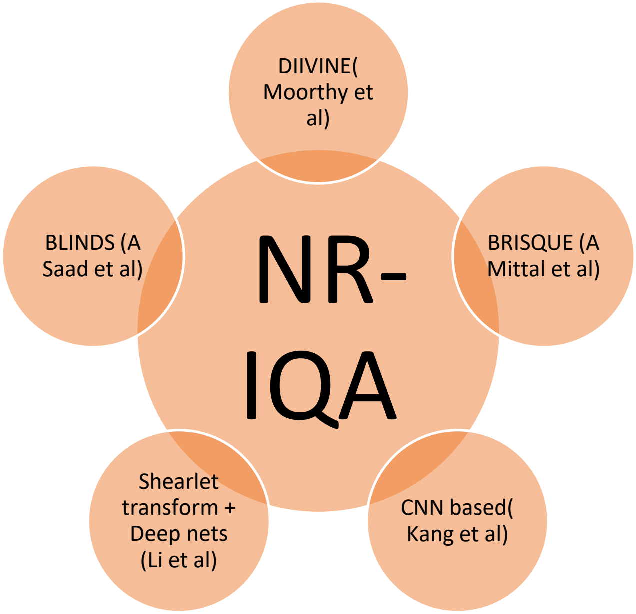
*by Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy and Megha Manohara*

At Netflix we care about video quality, and we care about measuring video quality accurately at scale. Our method, Video Multimethod Assessment Fusion (VMAF), seeks to reflect the viewer's perception of our streaming quality. We are open-sourcing this tool anc invite the research community to collaborate with us on this important project.

## Our Quest for High Quality Video

# No-reference image quality assessment

Data* → Features → Learning model → Perceptual quality



# Performance of existing algorithms on HDR data

| Scheme | Processing | SRCC | KRCC | PLCC | RMSE |
|---|---|---|---|---|---|
| **BRISQUE** | Lin | 0.5400 | 0.3732 | 0.4772 | 28.8475 |
| | PU | 0.7135 | 0.5121 | 0.6503 | 20.5534 |
| | TMO - Drago | 0.6337 | 0.4483 | 0.5903 | 21.7118 |
| | TMO - Reinhard 02 | 0.6583 | 0.4670 | 0.6512 | 18.4500 |
| | TMO - Reinhard 05 | 0.3524 | 0.2482 | 0.3946 | 30.6615 |
| | TMO - Mantiuk | 0.5887 | 0.4103 | 0.5493 | 22.7529 |
| **SSEQ** | Lin | 0.5287 | 0.3599 | 0.4714 | 25.2588 |
| | PU | 0.6492 | 0.4543 | 0.6111 | 19.6977 |
| | TMO - Drago | 0.5865 | 0.3956 | 0.5634 | 22.6987 |
| | TMO - Reinhard 02 | 0.5810 | 0.4075 | 0.5644 | 22.9900 |
| | TMO - Reinhard 05 | 0.4990 | 0.3401 | 0.5036 | 24.9193 |
| | TMO - Mantiuk | 0.4973 | 0.3398 | 0.4770 | 21.2044 |
| **BIQI** | Lin | 0.2845 | 0.1876 | 0.2831 | 31.0686 |
| | PU | 0.4386 | 0.3041 | 0.4399 | 21.2084 |
| | TMO - Drago | 0.5332 | 0.3780 | 0.4436 | 25.6200 |
| | TMO - Reinhard 02 | 0.4632 | 0.3196 | 0.4358 | 22.0376 |
| | TMO - Reinhard 05 | 0.5748 | 0.4048 | 0.5630 | 19.4825 |
| | TMO - Mantiuk | 0.4651 | 0.3204 | 0.4571 | 24.2268 |
| **DIIVINE** | Lin | 0.5041 | 0.3429 | 0.5209 | 20.6506 |
| | PU | 0.5318 | 0.3691 | 0.5442 | 19.6772 |
| | TMO - Drago | 0.4143 | 0.2852 | 0.4065 | 25.9697 |
| | TMO - Reinhard 02 | 0.3634 | 0.2434 | 0.3953 | 26.1464 |
| | TMO - Reinhard 05 | 0.5558 | 0.3849 | 0.5374 | 19.3122 |
| | TMO - Mantiuk | 0.4138 | 0.2838 | 0.4496 | 21.0499 |
| **kCNN** | Lin | 0.6991 | 0.5156 | 0.7008 | 19.3677 |
| **kCNN** | PU | 0.7694 | 0.5845 | 0.7544 | 18.5854 |

# PU experiment

- Tests on real images performed by [6].
  - 16 test subjects, assess the quality of the distorted image with respect to the reference on a 5 point scale.
  - 3 types of distortions (random pixel noise, gaussian blur and JPEG compression) at 2 levels (high and low) to 3 images.
  - Each image displayed on screens with contrast Low (1 – 100 Cd/m2) and High (10 – 1000Cd/m2)
  - Mean quality value for the Low brightness was 1.85, and for the High brightness was 2.15
- Low contrast display showed less errors – less sensitivity to errors.

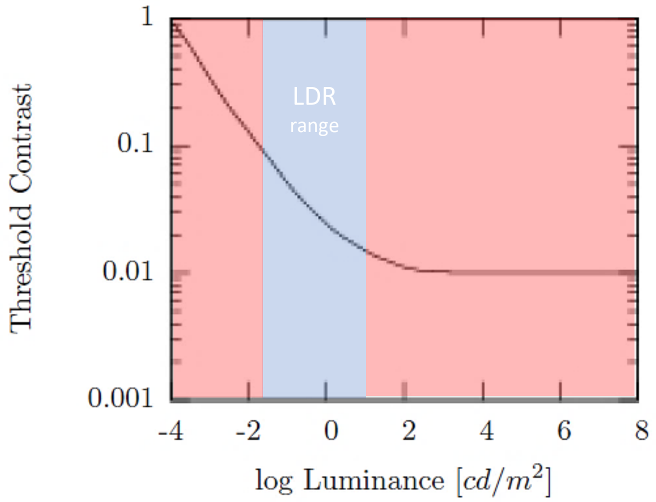
# Statistical assumptions in NR-IQA

- All the algorithms work by **assessing the noise statistics** and modelling its **impact on image quality**. Assumes Weber's Law holds.
- Not valid in HDR range.

$$\frac{\Delta I}{I} = k$$

**Threshold contrast** - Smallest change in intensity that is noticeable.

Intensity of the surroundings




Courtesy: Aydin et al. [6]

- **Need to Model additional perceptual effect in HDR.**

# Mathematical Framework: Quality Model

- For any distorted image patch with mean error $\delta(i, j)$ .
- The impact of the error $\delta(i, j)$ at location $(i, j)$ will be :

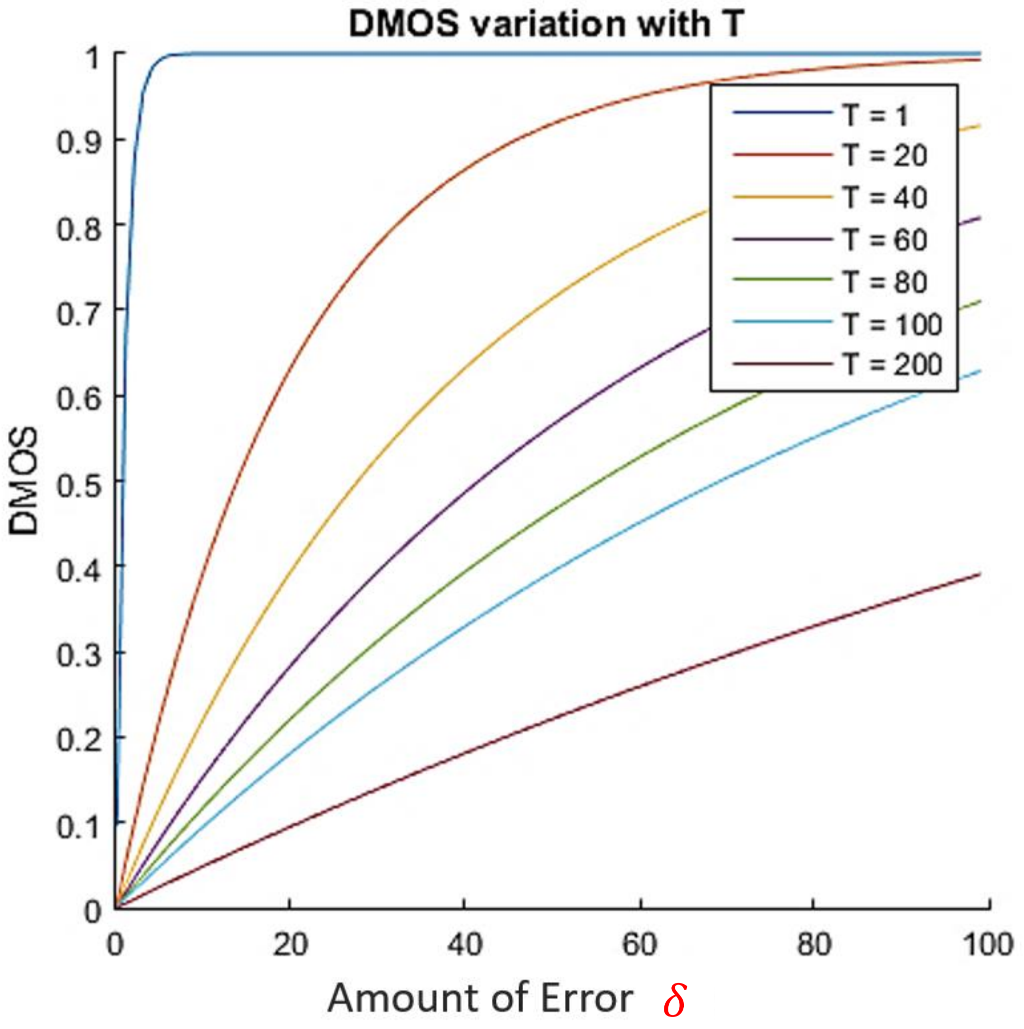
$$DMOS(i, j) = 1 - \exp(-|\frac{\delta(i,j)}{T(i,j)}|)$$

- Where $T(i, j)$ is the perceptual sensitivity threshold component for the image block.
- T() conventionally measured by pshycovisual experiments.

# Mathematical Framework

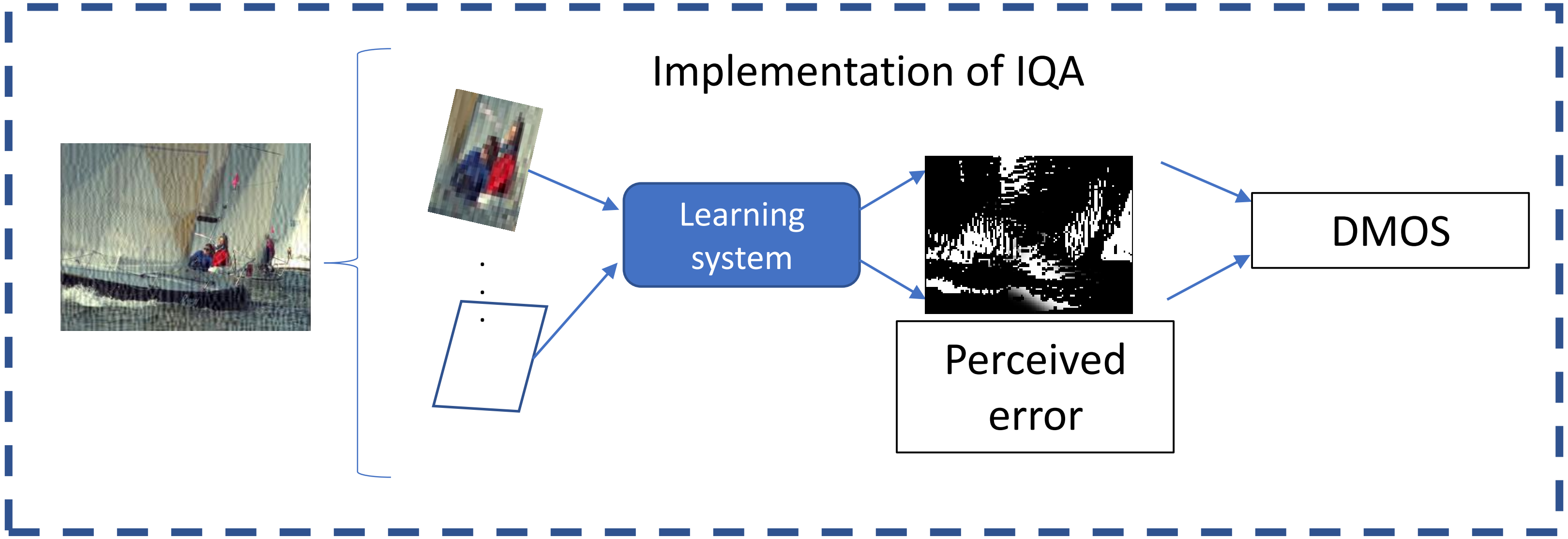
$$DMOS(i,j) = 1 - \exp(-|\frac{\delta(i,j)}{T(i,j)}|)$$

- DMOS ( perceived noise )
  - -> Increase with error
  - -> Decrease with sensitivity threshold



# Approach

- Implemented by using a **learning system** that predicts quality score of *every patch* of image.
- System trained with human quality score of image as target .

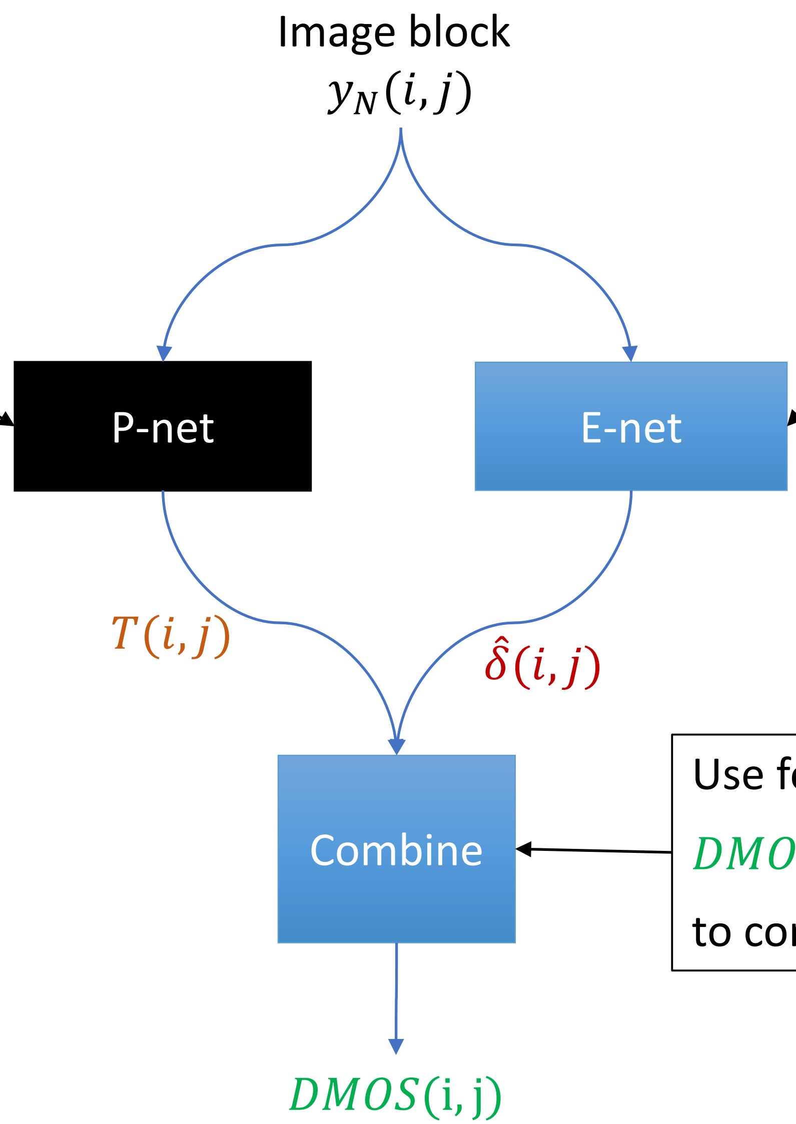



# Approach

- Process can be broken to two subcomponents with independent underlying distributions – Perceptual vs Statistical

- Model the **learning system** with **two *Convolutional Neural Nets***.
  - The first net learns to detects error - $\delta$.
  - The second net learns to detect perceptual thresholds – Perceptual Resistance - $T$.
  - Combine both these with quality model.

# Network structure




CNN that takes image block as input and gives a value $T()$ as output.

CNN that estimates error in the distorted image.

*T* is computed from data – NOT experiments with sine wave gratings.

Use formula

$$DMOS(i,j) = 1 - \exp(-|\frac{\hat{\delta}(i,j)}{T(i,j)}|)$$

to combine the two values

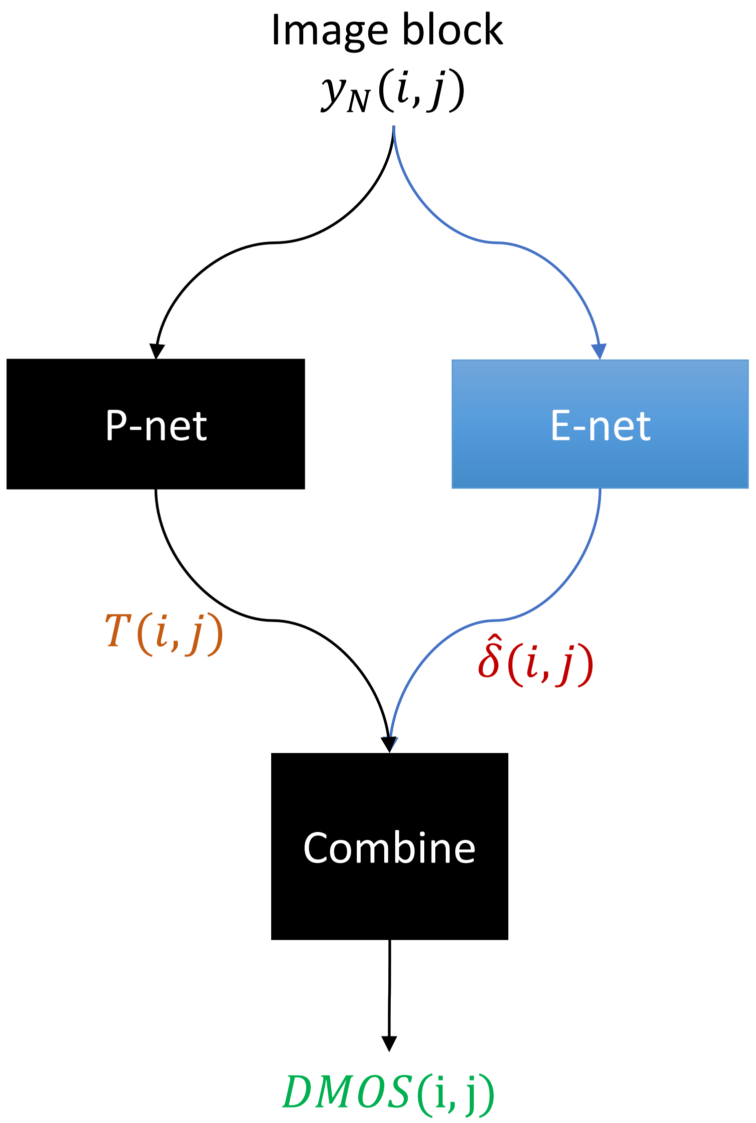
# Problems in Training

- Datasets :
  Clean-noisy image pair, DMOS for noisy image.
- E-net – easy to train.

  $\delta(i,j) = Y_N(\mathrm{i,j}) - Y_C(\mathrm{i,j})$

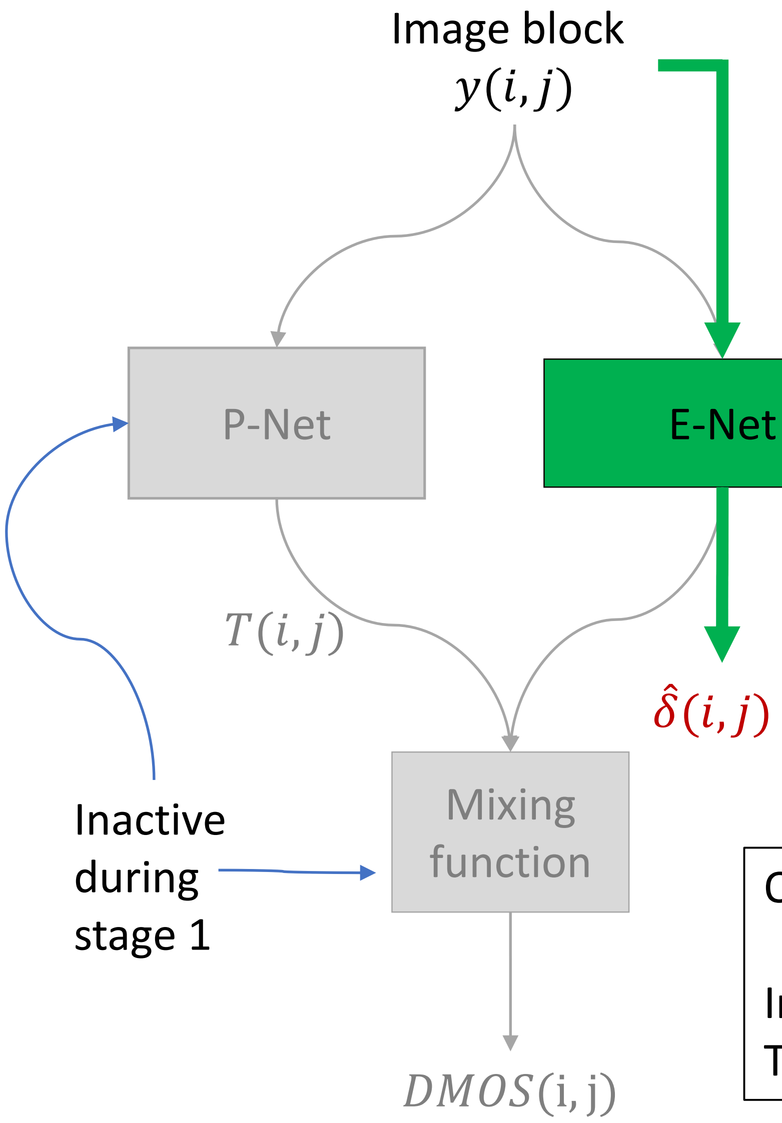
- P-net – train indirectly. Ground truth not known.

  Derive $T(i,j)$ that takes error $\delta(i,j)$ and produce a score equal to the ground truth DMOS using quality model.



# Network Training

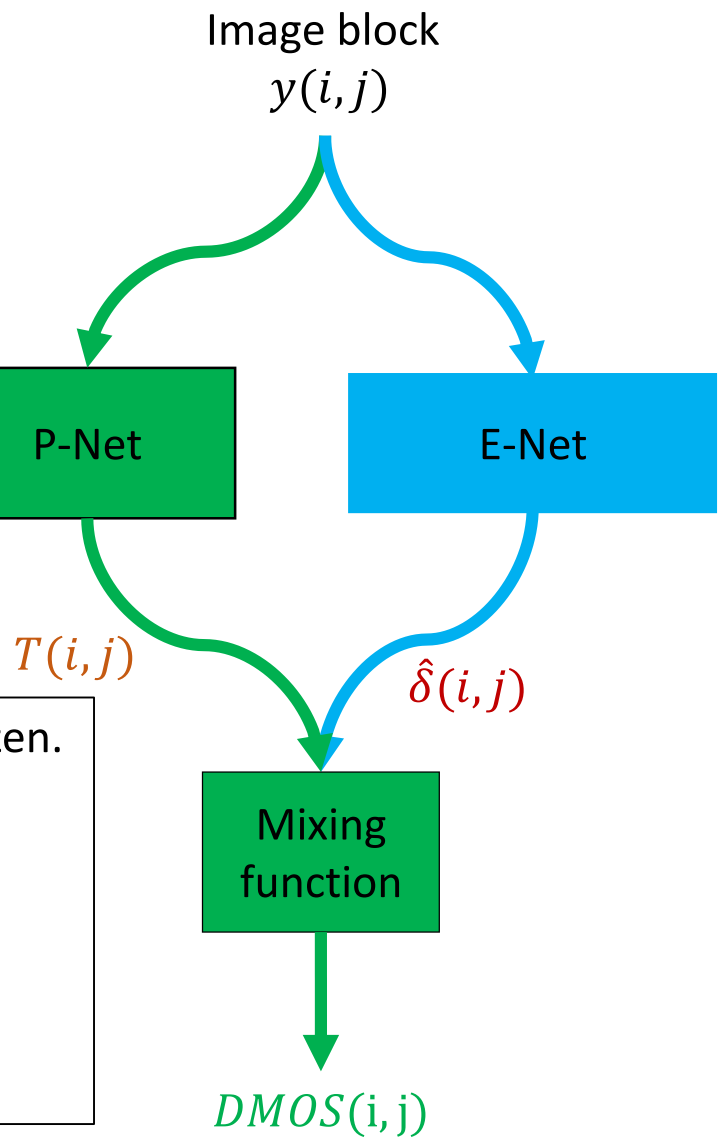
## Stage 1




Only E-Net is trained.

Input : <$y(i,j)$>
Target : $\delta(i,j)$.

## Stage 2




E-Net weights are frozen. P-Net and mixing function are trained together.

Input : <$y(i,j)$>
Target : $DMOS(i,j)$.

# Datasets

- Multiple datasets for HDR combined with iterated nested least square algorithm (INLSA) proposed by Pinson and Wolf.

| Dataset Number | Number of Reference Images | Number of Distorted Images | Distortion type | Maximum Luminance ($Cd/m^2$) |
|---|---|---|---|---|
| #1 | 27 | 140 | JPEG | 1000 |
| #2 | 29 | 210 | JPEG 2000 | 1000 |
| #3 | 24 | 240 | JPEG-XT | 4000 |
| #4 | 15 | 50 | JPEG<br>JPEG2000<br>JPEGXT | 4000 |
| #5 | 15 | 50 | JPEG<br>JPEG2000<br>JPEGXT | 4000 |

# Results on the full dataset

- Mean results of 100 cycles of train-test shown.

| Scheme | Processing | SRCC | KRCC | PLCC | RMSE |
|---|---|---|---|---|---|
| **BRISQUE** | Lin | 0.7274 | 0.5430 | 0.7231 | 18.1797 |
| | PU | 0.8047 | 0.6116 | 0.7825 | 17.3576 |
| | TMO - Drago | 0.7374 | 0.5415 | 0.7203 | 19.1261 |
| | TMO - Reinhard 02 | 0.7782 | 0.5853 | 0.7699 | 18.1523 |
| | TMO - Reinhard 05 | 0.6903 | 0.5061 | 0.6643 | 20.3307 |
| | TMO - Mantiuk | 0.6172 | 0.4559 | 0.6148 | 22.1868 |
| **SSEQ** | Lin | 0.6022 | 0.4378 | 0.6008 | 23.3017 |
| | PU | 0.7342 | 0.5451 | 0.7175 | 19.4117 |
| | TMO - Drago | 0.6853 | 0.5011 | 0.6954 | 20.8766 |
| | TMO - Reinhard 02 | 0.6866 | 0.5183 | 0.6688 | 21.0673 |
| | TMO - Reinhard 05 | 0.6568 | 0.4845 | 0.6467 | 20.5737 |
| | TMO - Mantiuk | 0.4185 | 0.2926 | 0.4651 | 25.7570 |
| **BIQI** | Lin | 0.1817 | 0.1391 | 0.1466 | 38.7513 |
| | PU | 0.3387 | 0.2406 | 0.3445 | 30.5220 |
| | TMO - Drago | 0.2803 | 0.1923 | 0.2960 | 41.0579 |
| | TMO - Reinhard 02 | 0.3756 | 0.2778 | 0.3766 | 33.2005 |
| | TMO - Reinhard 05 | 0.3097 | 0.2213 | 0.2874 | 27.7294 |
| | TMO - Mantiuk | 0.2822 | 0.1957 | 0.2408 | 39.0999 |
| **DIIVINE** | Lin | 0.6677 | 0.4853 | 0.6759 | 21.8020 |
| | PU | 0.7156 | 0.5290 | 0.7193 | 18.7586 |
| | TMO - Drago | 0.7418 | 0.5562 | 0.7400 | 18.9959 |
| | TMO - Reinhard 02 | 0.7149 | 0.5266 | 0.7024 | 20.7177 |
| | TMO - Reinhard 05 | 0.7900 | 0.5932 | 0.7809 | 17.2134 |
| | TMO - Mantiuk | 0.4946 | 0.3549 | 0.4936 | 27.4918 |
| **kCNN** | Lin | 0.8363 | 0.6530 | 0.8134 | 19.1753 |
| | PU | **0.8638** | **0.6852** | **0.8497** | **16.8937** |
| | TMO - Drago | 0.7700 | 0.5853 | 0.7485 | 18.2759 |
| | TMO - Mantiuk | 0.8075 | 0.6188 | 0.8053 | 17.7948 |
| | TMO - Reinhard 02 | 0.8613 | 0.6668 | 0.8179 | 17.7157 |
| | TMO - Reinhard 05 | 0.6438 | 0.4631 | 0.6074 | 22.3484 |
| **Proposed** | PU | **0.8860** | **0.7170** | **0.8871** | **16.4171** |
| **Proposed** | Lin | **0.8920** | **0.7184** | **0.8860** | **14.1464** |

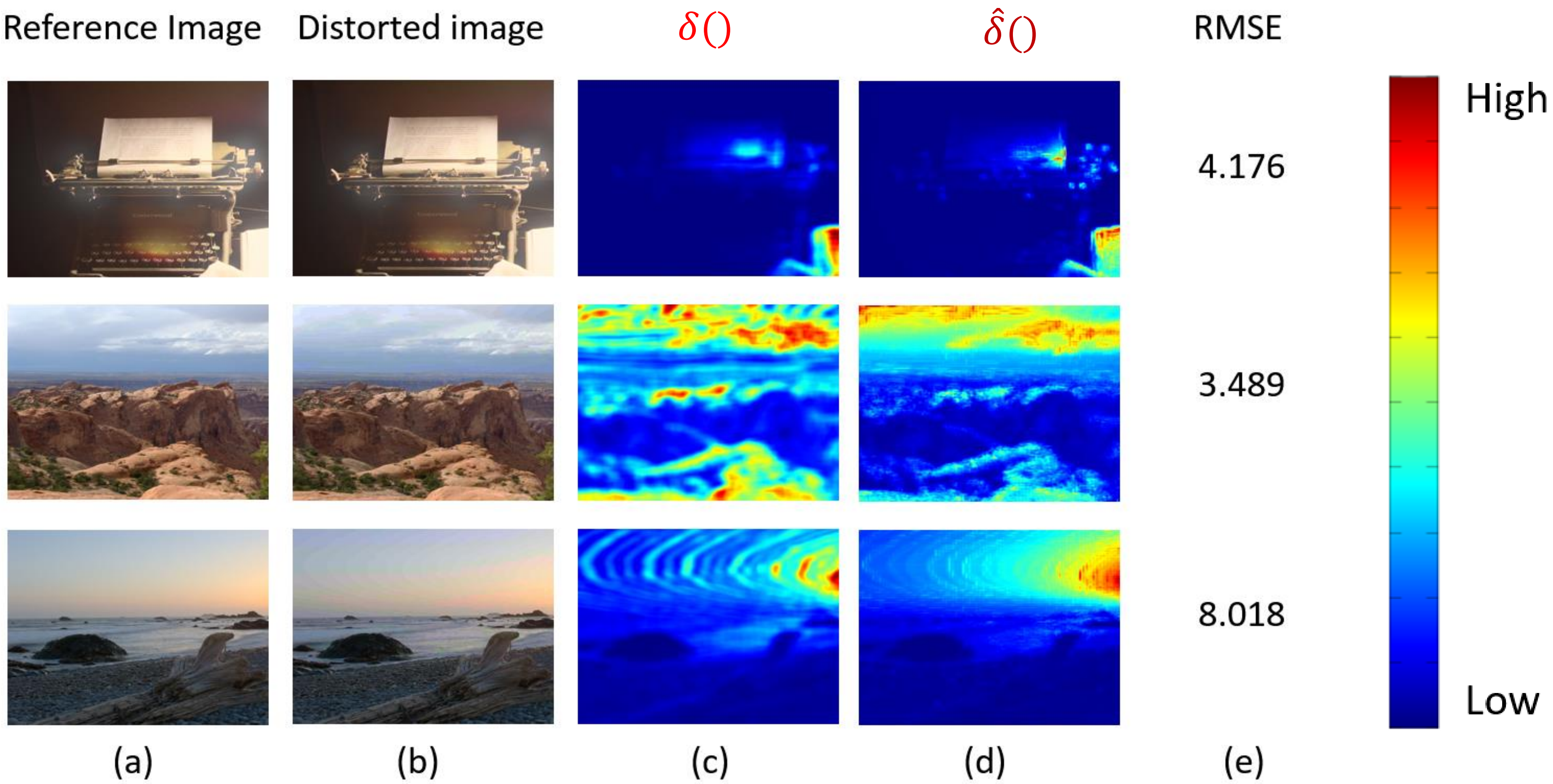
# Results on the full dataset

- Training on dataset [1] to [3] test on [4] and [5].

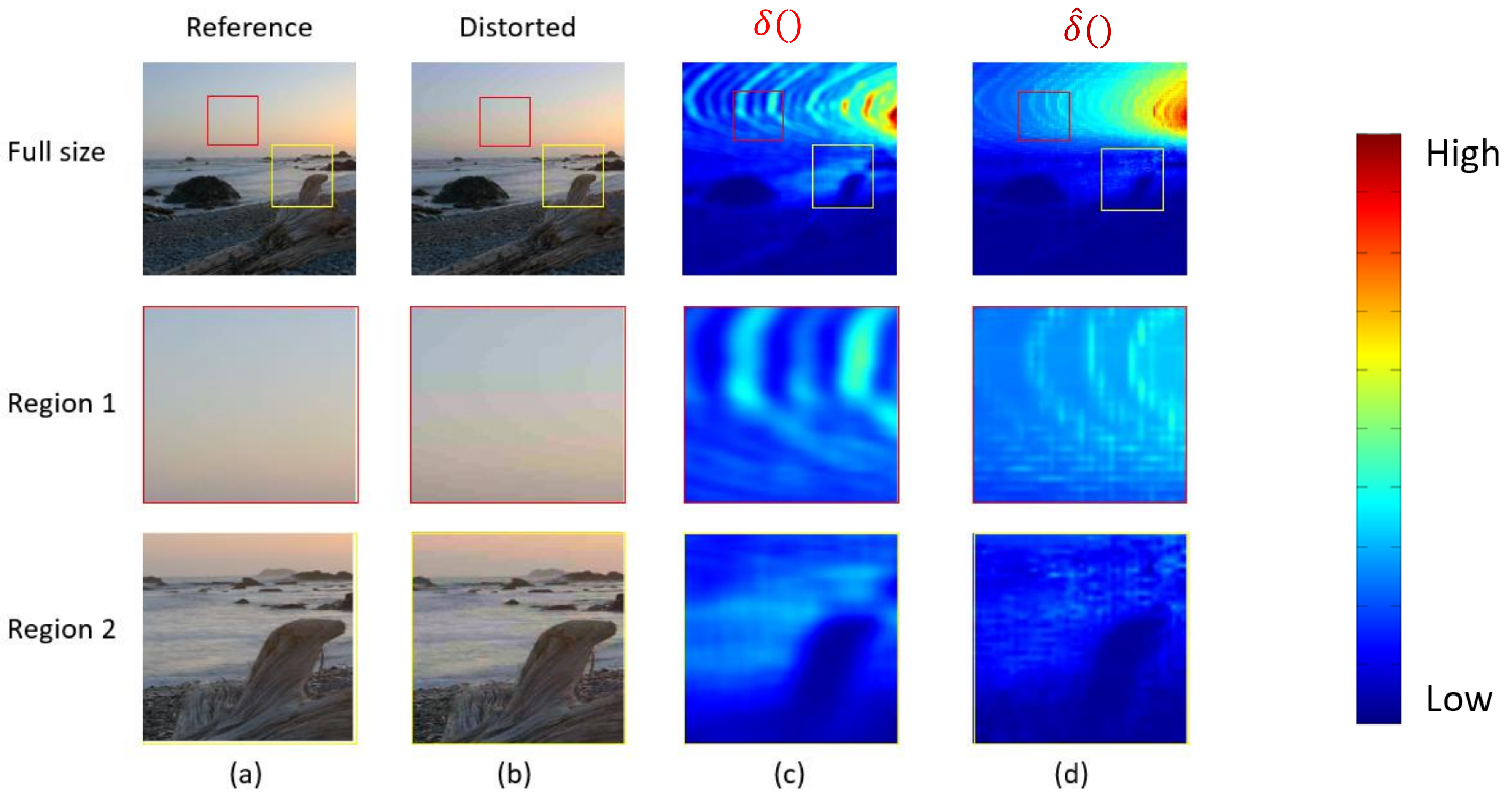
| Dataset Number | Number of Reference Images | Number of Distorted Images | Distortion type | Maximum Luminance ($Cd/m^2$) |
|---|---|---|---|---|
| #1 | 27 | 140 | JPEG | 1000 |
| #2 | 29 | 210 | JPEG 2000 | 1000 |
| #3 | 24 | 240 | JPEG-XT | 4000 |
| #4 | 15 | 50 | JPEG<br>JPEG2000<br>JPEGXT | 4000 |
| #5 | 15 | 50 | JPEG<br>JPEG2000<br>JPEGXT | 4000 |

| Scheme | Processing | SRCC | KRCC | PLCC | RMSE |
|---|---|---|---|---|---|
| BRISQUE | Lin | 0.5400 | 0.3732 | 0.4772 | 28.8475 |
| | PU | 0.7135 | 0.5121 | 0.6503 | 20.5534 |
| | TMO - Drago | 0.6337 | 0.4483 | 0.5903 | 21.7118 |
| | TMO - Reinhard 02 | 0.6583 | 0.4670 | 0.6512 | 18.4500 |
| | TMO - Reinhard 05 | 0.3524 | 0.2482 | 0.3946 | 30.6615 |
| | TMO - Mantiuk | 0.5887 | 0.4103 | 0.5493 | 22.7529 |
| SSEQ | Lin | 0.5287 | 0.3599 | 0.4714 | 25.2588 |
| | PU | 0.6492 | 0.4543 | 0.6111 | 19.6977 |
| | TMO - Drago | 0.5865 | 0.3956 | 0.5634 | 22.6987 |
| | TMO - Reinhard 02 | 0.5810 | 0.4075 | 0.5644 | 22.9900 |
| | TMO - Reinhard 05 | 0.4990 | 0.3401 | 0.5036 | 24.9193 |
| | TMO - Mantiuk | 0.4973 | 0.3398 | 0.4770 | 21.2044 |
| BIQI | Lin | 0.2845 | 0.1876 | 0.2831 | 31.0686 |
| | PU | 0.4386 | 0.3041 | 0.4399 | 21.2084 |
| | TMO - Drago | 0.5332 | 0.3780 | 0.4436 | 25.6200 |
| | TMO - Reinhard 02 | 0.4632 | 0.3196 | 0.4358 | 22.0376 |
| | TMO - Reinhard 05 | 0.5748 | 0.4048 | 0.5630 | 19.4825 |
| | TMO - Mantiuk | 0.4651 | 0.3204 | 0.4571 | 24.2268 |
| DIIVINE | Lin | 0.5041 | 0.3429 | 0.5209 | 20.6506 |
| | PU | 0.5318 | 0.3691 | 0.5442 | 19.6772 |
| | TMO - Drago | 0.4143 | 0.2852 | 0.4065 | 25.9697 |
| | TMO - Reinhard 02 | 0.3634 | 0.2434 | 0.3953 | 26.1464 |
| | TMO - Reinhard 05 | 0.5558 | 0.3849 | 0.5374 | 19.3122 |
| | TMO - Mantiuk | 0.4138 | 0.2838 | 0.4496 | 21.0499 |
| **kCNN** | Lin | 0.6991 | 0.5156 | 0.7008 | 19.3677 |
| **kCNN** | PU | 0.7694 | 0.5845 | 0.7544 | 18.5854 |
| **Proposed** | Lin | **0.8672** | **0.6773** | **0.8780** | **18.626** |

# Results: E-net

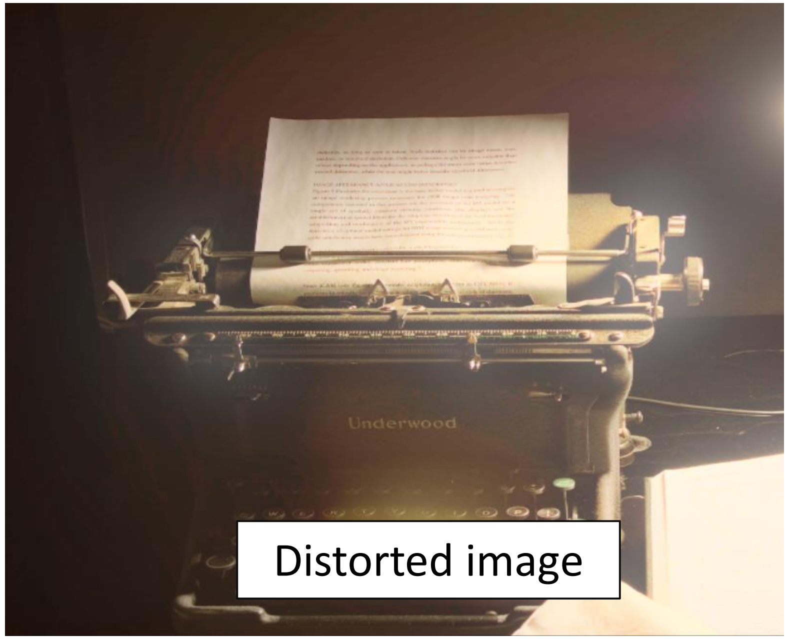


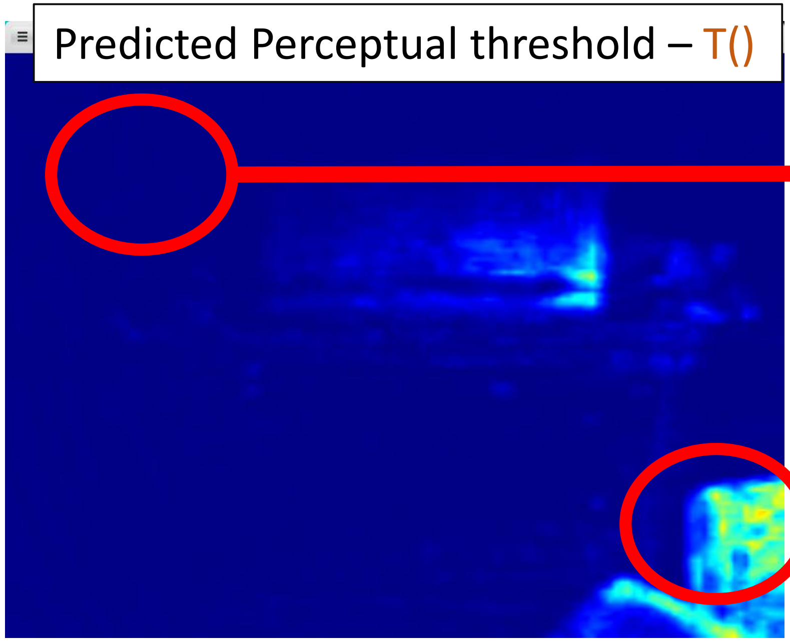
# Results: E-net




CNN's good at estimating noise.

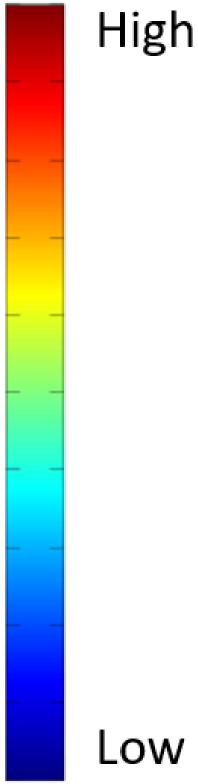
# Results: P-net




Low thresholds indicate that the errors here can easily detected

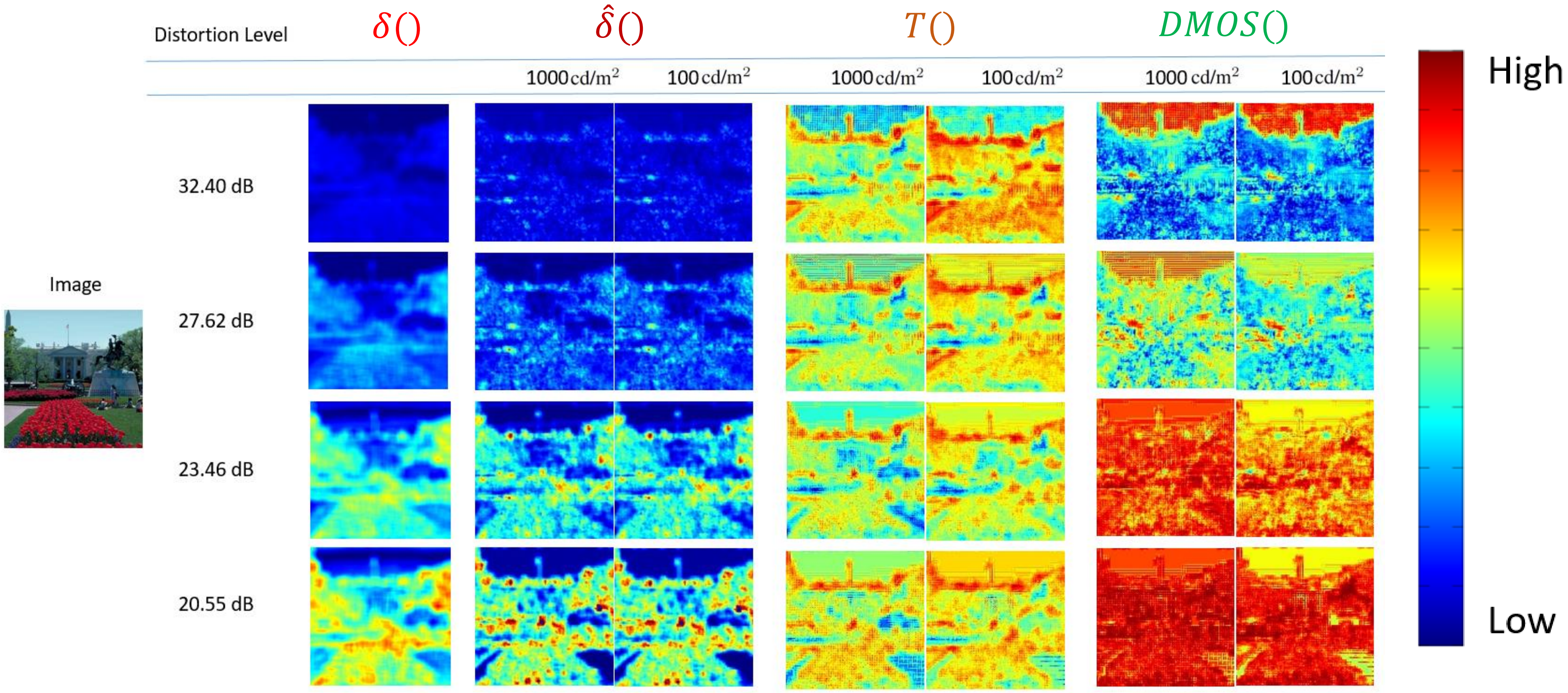
High detection thresholds at extremely bright regions.



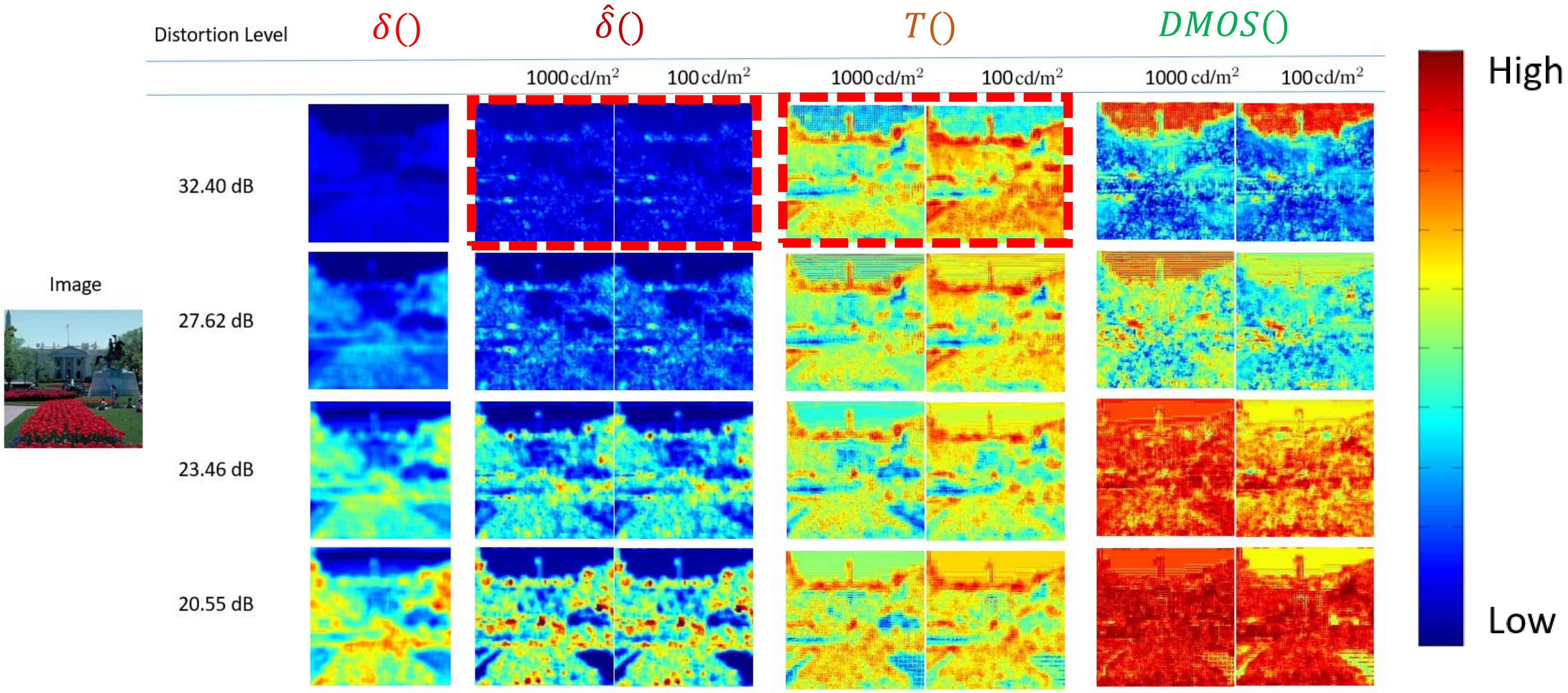
# Results: P-net (Replication of PU experiment)

- We used JPEG distorted images - ***scaled the images to luminance 100 cd/m2 and 1,000 cd/m2***, respectively - ***computed the quality*** *of these images* using the proposed method.
- We found that the mean value of DMOS for the images of low brightness was 68.76 and those of high brightness was 76.72.
- Difference found to be is statistically significant.
- Same results as the original paper – data driven method can recover perceptual thresholds.

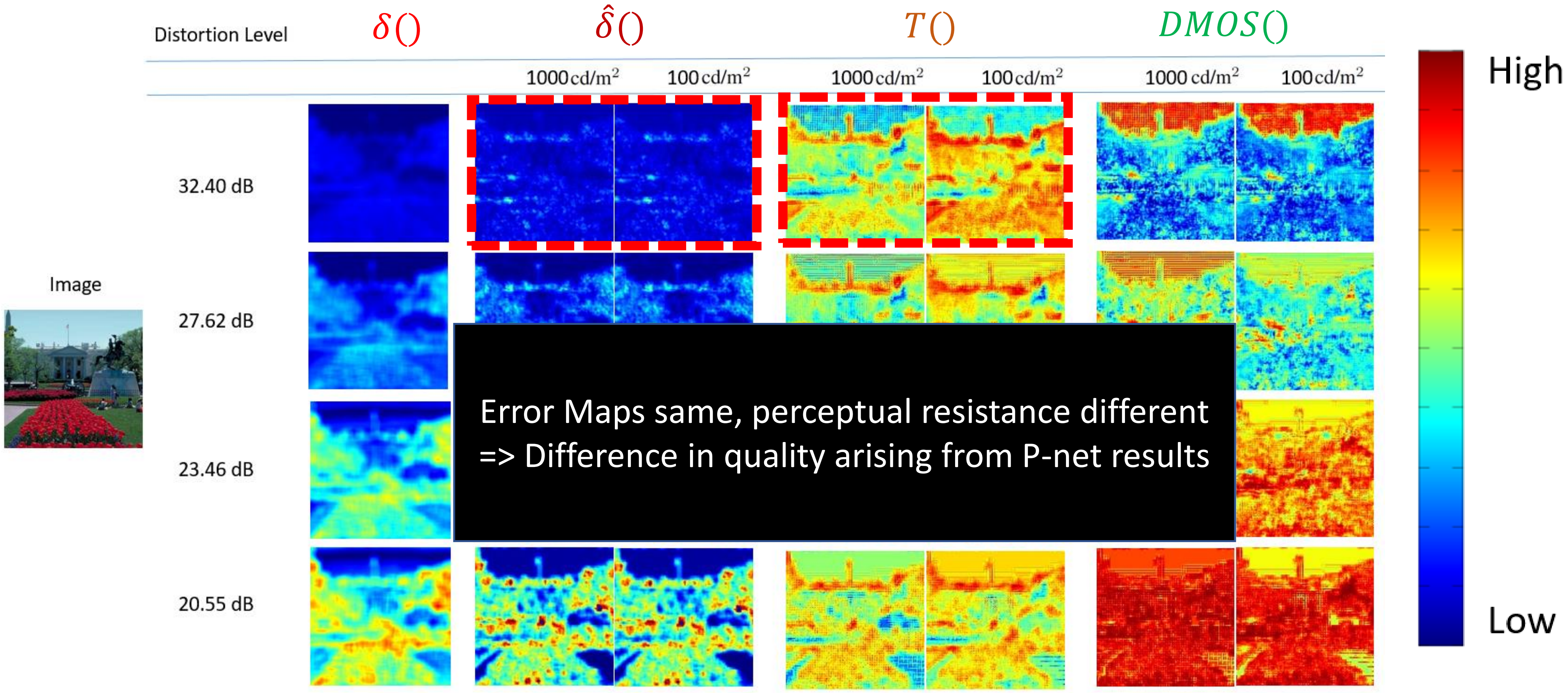
# Results: P-net



# Results: P-net



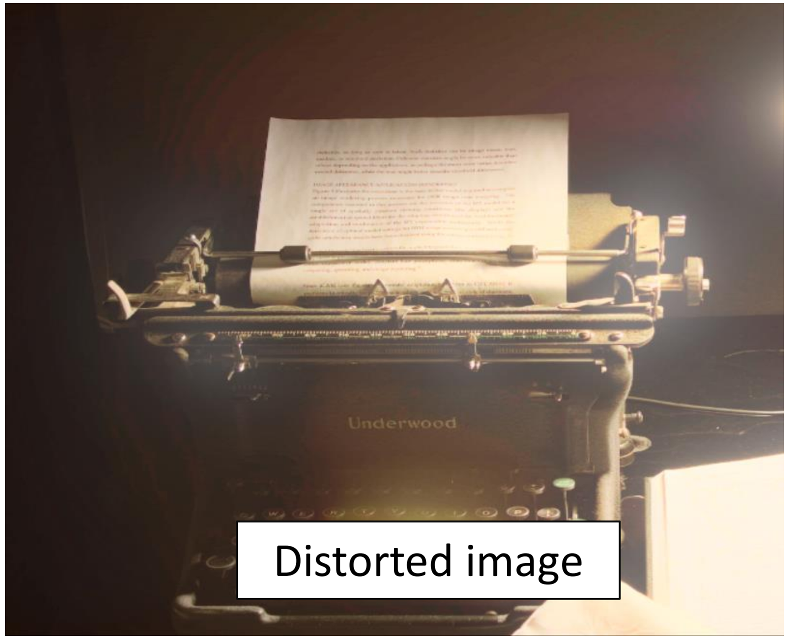
# Results: P-net

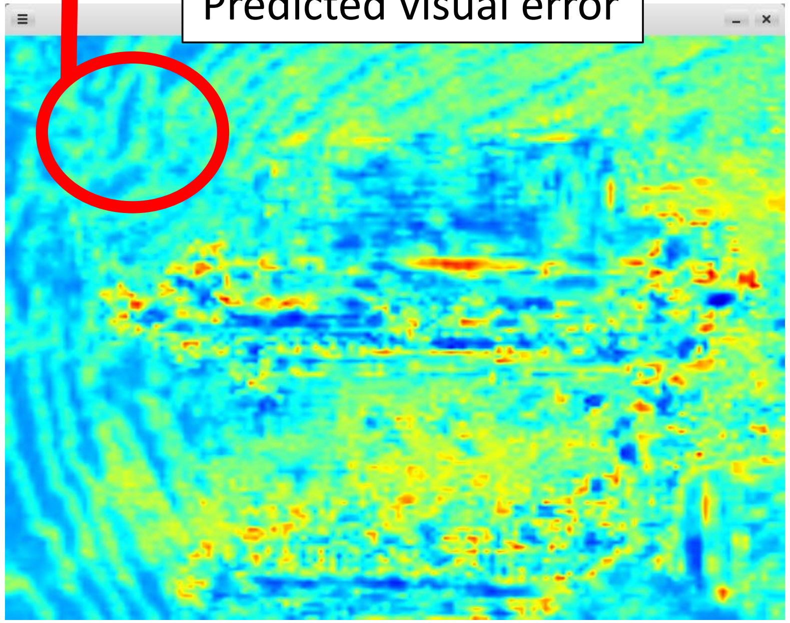



Error Maps same, perceptual resistance different
=> Difference in quality arising from P-net results

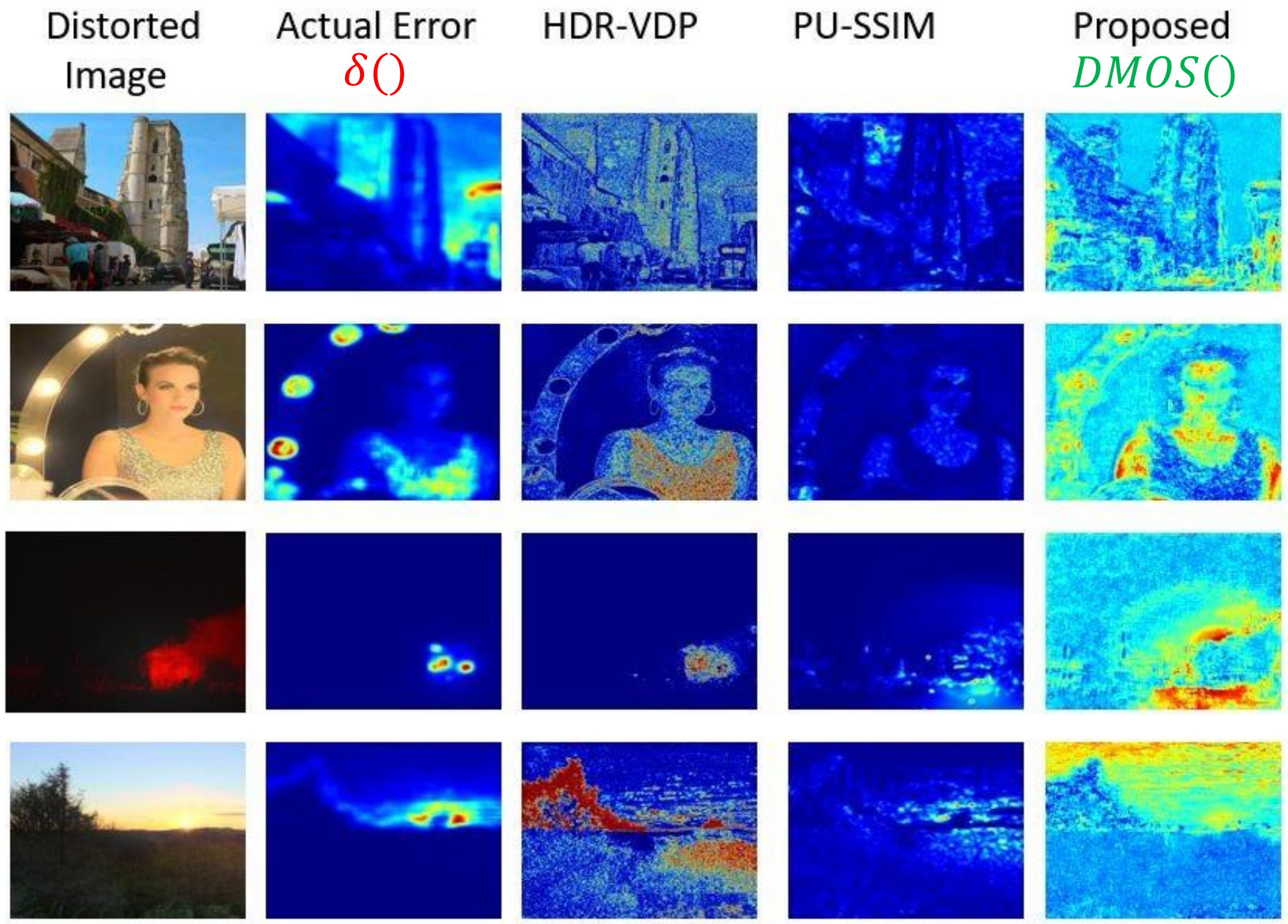
# Results: Error Map

Detects banding in smooth backgrounds.



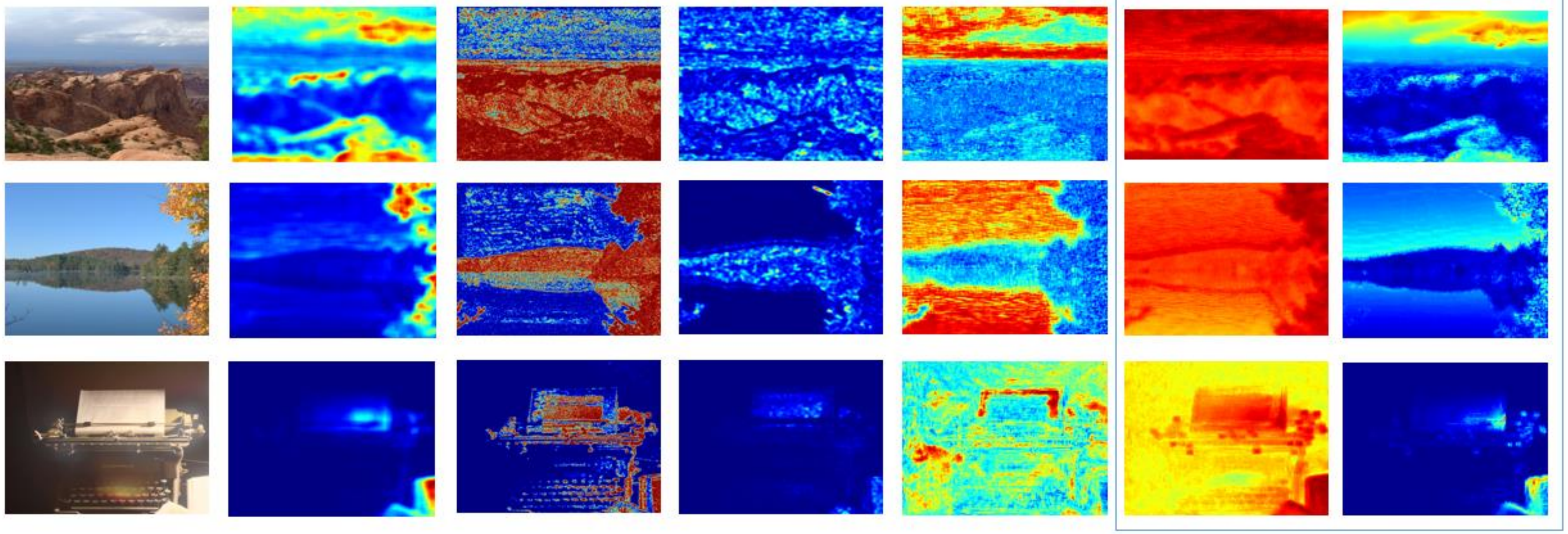



# Error map



Distorted Image
Actual Error
$\delta()$
HDR-VDP
PU-SSIM
Proposed
$DMOS()$

# Failure cases



| Distorted Image | $\delta$ | HDR-VDP 2.2 | PU-SSIM | Perceptual Distortion Map | $T$ | $\hat{\delta}$ | MSE with HDR-VDP | DMOS Actual | DMOS Predicted |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 0.1833 | 29 | 44 |
| | | | | | | | 0.2627 | 37 | 57 |
| | | | | | | | 0.2361 | 42 | 61 |

# Conclusion

- Learning based system that can predict quality of compressed HDR images.
- Can estimate perceived errors and error detection thresholds.
- Produces state of the art results in terms of similarity to human judgement.

# Team

Dr Giuseppe Valenzise

Dr Frederic Dufaux

Navaneeth K. K.

Dr Irene Cheng

# Thanks

Emin Zerman

Dr Anup Basu

# References

[1] M. Narwaria, M. P. Da Silva, P. Le Callet, and R. Pepion, “*Tone mapping-based high-dynamic-range image compression: study of optimization criterion and perceptual quality*,” Optical Engineering, vol. 52, no. 10, pp. 102 008–102 008, 2013.

[2] M. Narwaria, M. P. Da Silva, P. Le Callet, and R. P´epion, “*Impact of tone mapping in high dynamic range image compression*,” in International Workshop on Video Processing and Quality Metrics for Consumer Electronics, 2014, pp. pp–1.

[3] P. Korshunov, P. Hanhart, T. Richter, A. Artusi, R. Mantiuk, and T. Ebrahimi, “*Subjective quality assessment database of HDR images compressed with JPEG XT*,” in Quality of Multimedia Experience (QoMEX), 2015 Seventh International Workshop on. IEEE, 2015, pp. 1–6.

[4] G. Valenzise, F. De Simone, P. Lauga, and F. Dufaux, “*Performance evaluation of objective quality metrics for HDR image compression*,” in SPIE Optical Engineering Applications. International Society for Optics and Photonics, 2014, pp. 92 170C–92 170C.

[5] E. Zerman, G. Valenzise, and F. Dufaux, “*An extensive performance evaluation of full-reference HDR image quality metrics*.” in Springer: Quality and User Experience, vol. 2, no. 1, 2017.

[6] Aydin, Tunç Ozan, et al. "Dynamic range independent image quality assessment." *ACM Transactions on Graphics (TOG)*. Vol. 27. No. 3. ACM, 2008.


Thank
you